

**PENYEJAJARAN LOKAL BARISAN DNA DENGAN
MENGUNAKAN
METODE SMITH WATERMAN**

ILHAM CANDRA BUDIMAN

0303010214



**DEPARTEMEN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS INDONESIA
2009**

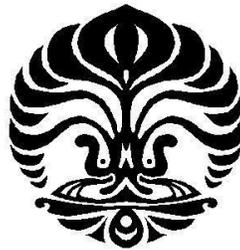
**PENYEJAJARAN LOKAL BARISAN DNA DENGAN MENGGUNAKAN
METODE SMITH WATERMAN**

**Skripsi diajukan sebagai salah satu syarat
untuk memperoleh gelar Sarjana Sains**

Oleh:

ILHAM CANDRA BUDIMAN

0303010214



DEPOK

2009

SKRIPSI : PENYEJAJARAN LOKAL BARISAN DNA DENGAN MENGGUNAKAN
METODE SMITH-WATERMAN

NAMA : ILHAM CANDRA BUDIMAN

NPM : 0303010214

SKRIPSI INI TELAH DIPERIKSA DAN DISETUJUI

DEPOK, 6 Juli 2009

Prof. Dr. Djati Kerami
Pembimbing I

Helen Burhan S.Si., M.Si
Pembimbing II

Tanggal lulus Ujian Sidang Sarjana : 6 Juli 2009

Penguji I : Prof. Dr. Djati Kerami

Penguji II : Dra. Siti Aminah, M.kom.

Penguji III : Alhaji Akbar B.,S.Si.,M.Sc.

SKRIPSI : PENYEJAJARAN LOKAL BARISAN DNA DENGAN MENGGUNAKAN METODE SMITH-WATERMAN

NAMA : ILHAM CANDRA BUDIMAN

NPM : 0303010214

SKRIPSI INI TELAH DIPERIKSA DAN DISETUJUI

DEPOK, 6 JULI 2009

Prof. Dr. Djati Kerami
PEMBIMBING I

Helen Burhan S.si ., M.si
PEMBIMBING II

Tanggal lulus Ujian Sidang Sarjana : 6 Juli 2009

Penguji I : Prof. Dr. Djati Kerami

Penguji II : Dra. Siti Aminah, M.kom.

Penguji III : Alhaji Akbar B.,S.Si.,M.Sc.

KATA PENGANTAR

Segala puji bagi Allah SWT yang telah memberikan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi ini. Shalawat dan salam semoga tercurah kepada junjungan Rasulullah Muhammad SAW beserta keluarga, sahabat, dan para pengikutnya hingga akhir zaman.

Penulis menyadari bahwa skripsi ini tidak akan selesai tanpa bantuan, dorongan, dan do'a dari orang-orang di sekitar penulis. Oleh karena itu penulis ingin mengucapkan terima kasih kepada semua pihak yang telah membantu penulis dalam menyelesaikan skripsi ini, khususnya kepada:

1. Prof. Dr. Djati Kerami, selaku pembimbing I, terima kasih atas segala saran dan bimbingannya selama ini serta penulis ingin meminta maaf atas segala kesalahan yang dilakukan oleh penulis baik yang disengaja maupun tidak disengaja.
2. Ibu Helen Burhan S.si.,M.si, selaku pembimbing II, terima kasih atas segala saran dan bimbingannya selama ini.
3. Ibu Bevina, selaku pembimbing akademik penulis, terima kasih atas saran, bimbingan, dan dorongan semangat selama penulis menempuh perkuliahan di matematika.

4. Mama, Papa serta kak Wahyu terima kasih atas segala dukungan dan pengorbanannya untuk menjadikan penulis seperti sekarang ini. Beribu-ribu cinta dari penulis tidak akan cukup untuk membalas segala pengorbanan yang kalian berikan.
5. Seluruh dosen matematika UI yang tidak bisa disebutkan satu-persatu. Terima kasih atas bimbingannya sehingga penulis memperoleh pengalaman akan luasnya dunia matematika.
6. Gilang, Gunung, Bembi, Yessamah, Diki, Soni, dan Igun yang telah membantu penulis dalam menyelesaikan tugas akhir ini.
7. Dodi, Toni, Arief, dan Adri, semoga kalian berhasil mencapai tujuan kalian.
8. Semua teman-teman di matematika UI yang telah membantu penulis dalam proses penyelesaian skripsi ini.
9. Seluruh karyawan matematika UI terutama mba Santi, terima kasih atas bantuannya selama penulis kuliah di matematika UI.

Penulis menyadari bahwa skripsi ini jauh dari sempurna, oleh karena itu kritik dan saran yang membangun sangat diharapkan. Semoga skripsi ini berguna bagi penelitian selanjutnya.

Penulis

2009

ABSTRAK

Penyejajaran antar barisan DNA digunakan untuk melihat tingkat kemiripan dari barisan DNA tersebut. Sebagian besar metode dalam penyejajaran barisan menggunakan pendekatan program dinamik. Salah satu metode yang sering digunakan adalah metode Needleman-Wunsch. Pada metode Needleman-Wunsch semua karakter pada barisan-barisan tersebut disejajarkan sehingga dapat terlihat kemiripan dari barisan-barisan DNA tersebut.

Metode yang digunakan dalam tugas akhir ini tidak menyejajarkan seluruh karakter dari barisan-barisan DNA. Metode ini hanya menyejajarkan pasangan segmen dari dua barisan DNA. Hasil dari metode ini adalah pasangan segmen dari dua barisan DNA yang memiliki kemiripan paling besar.

Kata kunci : Penyejajaran Lokal, Smith-Waterman, Penyejajaran Dua Barisan

v+52 hlm.;lamp

Bibliografi: 3 (1981 - 2006)

DAFTAR ISI

KATA PENGANTAR	i
ABSTRAK	iii
DAFTAR ISI.....	iv
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Perumusan Masalah	3
1.3. Tujuan Penulisan.....	4
1.4. Pembatasan Masalah.....	4
1.5. Sistematika Penulisan.....	4
BAB II PENYEJAJARAN BARISAN	5
2.1. Definisi Penyejajaran Dua Barisan	5
2.2. Pemberian Bobot pada Penyejajaran Dua Barisan DNA	8
2.3. Menentukan Penyejajaran yang Optimum	10
2.4. Penyejajaran Global dan Penyejajaran Lokal.....	12
BAB III METODE SMITH-WATERMAN.....	14
3.1. Menentukan Pembobotan pada Metode Smith-Waterman	15
3.2. Inisialisasi Pada Metode Smith-Waterman.....	18
3.3. Penelusuran Kembali.....	23
3.3. Contoh Penyejajaran Smith-Waterman	24
3.4. Penggunaan Progam Dinamik pada Metode Smith-Waterman.....	36

BAB IV IMPLEMENTASI.....	41
4.1. Hasil Implementasi	41
BAB V KESIMPULAN	48
DAFTAR PUSTAKA.....	49
LAMPIRAN.....	.50



BAB I

PENDAHULUAN

1.1 Latar Belakang

Bioinformatika adalah ilmu yang mempelajari penerapan teknik komputasional untuk mengelola dan menganalisis informasi biologis. Bidang ini mencakup penerapan metode-metode matematika, statistika, dan informatika untuk memecahkan masalah-masalah biologis, terutama yang berhubungan dengan barisan DNA, asam amino serta informasi yang berkaitan dengannya.

Salah satu masalah yang cukup sering dibahas adalah barisan DNA (Deoxyribonucleid Acid). DNA merupakan polimer yang terdiri dari tiga komponen utama, yaitu gugus fosfat, gula deoksiribosa, dan basa nitrogen. DNA terdiri atas dua untai yang berpilin membentuk struktur heliks ganda. Pada struktur heliks ganda, orientasi rantai nukleotida pada satu untai berlawanan dengan orientasi nukleotida untai lainnya. Hal ini disebut sebagai *antiparalel*. Masing-masing untai terdiri dari rangka utama, sebagai struktur utama, dan basa nitrogen, yang berinteraksi dengan untai DNA satunya pada heliks. Kedua untai pada heliks ganda DNA disatukan oleh ikatan hidrogen antara basa-basa yang terdapat pada kedua untai tersebut. Empat basa yang ditemukan pada DNA adalah adenine (dilambangkan A), sitosin (C, dari *cytosine*), guanin (G), dan timin

(T). Adenin berikatan hidrogen dengan timin, sedangkan guanin berikatan dengan sitosin. Keempat basa nitrogen tersebut membentuk suatu barisan nucleotida dimana setiap individu memiliki suatu barisan nucleotida yang unik. Suatu barisan DNA dapat dinyatakan sebagai suatu barisan karakter adenine (dilambangkan A), sitosin (C, dari *cytosine*), guanin (G), dan timin (T).

Dalam jangka waktu tertentu pada suatu barisan DNA dapat berevolusi sehingga terbentuk suatu barisan DNA baru yang memiliki kemiripan dengan DNA sebelumnya. Evolusi pada barisan DNA itu dapat berupa perubahan (substitution) karakter, penghilangan (deletion) karakter, dan penyisipan (insertion) karakter. Setiap organisme dalam suatu species memiliki kemiripan dalam batas tertentu.

Untuk melihat seberapa besar tingkat kemiripan dari dua barisan DNA dilakukanlah penyejajaran barisan. Penyejajaran barisan merupakan metode dasar dalam menganalisa barisan. Metode ini digunakan untuk mempelajari evolusi barisan-barisan DNA yang berasal dari leluhur yang sama.

Pada umumnya penyejajaran barisan menggunakan program dinamik. Beberapa metode yang sering digunakan untuk menyejajarkan barisan adalah Needleman-Wunsch. Pada metode Needleman-wunsch digunakan untuk menyusun penyejajaran global diantara dua atau lebih barisan.

Pada metode ini semua kemungkinan penyelarasan dipertimbangkan untuk memperoleh penyelarasan yang optimum. Tujuan dari penggunaan metode Needleman-Wuncsh adalah untuk melihat kemiripan dari dua buah barisan DNA atau lebih. Jika kemiripan dari dua barisan DNA ini tinggi maka dipercaya kedua barisan DNA ini berasal dari satu leluhur yang sama.

Dalam banyak aplikasi, dua barisan belum tentu mempunyai similaritas secara keseluruhan, tetapi bisa saja hanya mengandung pasangan segmen dari kedua barisan yang memiliki kemiripan yang tinggi. Metode yang biasa digunakan untuk mencari pasangan segmen tersebut adalah metode Smith-Waterman

Dalam tugas akhir ini kita akan mencari pasangan segmen dari dua barisan DNA yang telah disejajarkan yang memiliki tingkat kemiripan paling tinggi dari dari pasangan segmen lainnya menggunakan metode smith-Waterman.

1.2 Perumusan Masalah

Bagaimana menentukan pasangan segmen dari dua barisan yang memiliki kemiripan paling besar.

1.3 Tujuan Penulisan

Tujuan dari penulisan tugas akhir ini adalah ingin melakukan penyejajaran segmen dari dua barisan yang memiliki kemiripan paling tinggi dengan menggunakan metode Smith-Waterman.

1.4 Pembatasan Masalah

- Jumlah barisan yang ada dibatasi hanya untuk 2 barisan saja.

1.5 Sistematika Penulisan

Tugas akhir ini terdiri dari lima bab. Bab I berisi pendahuluan. Bab II berisi landasan teori yang membahas tentang Penyejajaran barisan. Bab III membahas metode Penyejajaran lokal pada dua barisan dengan menggunakan metode Smith-Waterman. Bab IV berisi implementasi metode Smith-Waterman dengan menggunakan software Matlab 7.0.1. Bab V berisi kesimpulan yang didapat dari hasil Penelitian ini.

BAB II

PENYEJAJARAN BARISAN

Dalam proses mencocokkan dua atau lebih DNA tidak mungkin kita menemukan kesamaan 100% karena seperti dijelaskan sebelumnya, DNA bersifat unik. Namun, karena panjang DNA (jumlah rangkaian gugus karbon) dua atau lebih organisme tidak selalu sama, maka proses yang digunakan bukanlah 'perbandingan', melainkan 'penyejajaran'. Seperti dijelaskan sebelumnya suatu barisan DNA dapat dinyatakan sebagai barisan karakter dari adenine (dilambangkan A), sitosin (C, dari *cytosine*), guanin (G), dan timin (T). Hal yang sama juga diterapkan dalam proses penyejajaran protein.

Penyejajaran DNA dapat dilakukan antar dua barisan (pairwise Alignment) dan lebih dari dua barisan (multiple Alignment). Dalam tugas akhir ini yang akan dibahas adalah konsep dasar dari penyejajaran segmen antara dua barisan DNA.

2.1. Definisi penyejajaran dua barisan

Penyejajaran dua buah DNA adalah memberikan nilai kecocokan kedua DNA tersebut. Semakin tinggi nilainya, berarti kedua DNA tersebut semakin mirip satu sama lain, dan sebaliknya semakin rendah nilainya, berarti kedua DNA tersebut semakin tidak mirip.

Sebelum membahas penyejajaran dua barisan DNA, kita akan berikan terlebih dahulu definisi sekaligus ketentuan suatu penyejajaran barisan.

Suatu penyejajaran barisan (*sequence alignment*) dapat didefinisikan dengan ketentuan sebagai berikut :

1. Karakter-karakter dari suatu barisan DNA dipasangkan dengan karakter-karakter dari barisan DNA lainnya
2. Adanya kemungkinan penyisipan spasi di depan, di belakang atau ditengah-tengah barisan karakter sehingga barisan-barisan tersebut memiliki panjang yang sama.
3. Urutan karakter pada barisan DNA yang telah disejajarkan sama dengan urutan karakter pada barisan semula.
4. Tidak terjadi pemasangan seluruh karakter berupa spasi pada barisan-barisan tersebut.

contoh 1 :

Diberikan dua barisan , yaitu :

Barisan $X = x_1x_2x_3\dots x_{n-1}x_n$

Barisan $Y = y_1y_2y_3\dots y_{m-1}y_m$

Dimana x_i menyatakan karakter ke i pada barisan X dan y_j menyatakan karakter ke j pada barisan Y . penyejajaran barisan X dan Y merupakan

pemasangan karakter-karakter pada barisan X dan barisan Y berdasarkan ketentuan penyejajaran barisan yang telah dijelaskan sebelumnya.

dari penjelasan tersebut terdapat kemungkinan pemasangan karakter untuk penyejajaran barisan X dan barisan Y yaitu :

1. Pemasangan karakter tanpa adanya spasi

$$\begin{array}{ccccccc} X_1 & X_2 & X_3 & \dots & X_i & \dots & X_n \\ Y_1 & Y_2 & Y_3 & \dots & Y_j & \dots & Y_m \end{array}$$

Dimana terdapat dua kemungkinan pada penyejajaran diatas yaitu kedua karakter sama ($x_i = y_j$) atau kedua karakter berbeda ($x_i \neq y_j$), dimana x_i dan y_j bukan spasi.

2. Pemasangan karakter dengan spasi

Bisa terjadi dua kemungkinan yaitu

a. Pemasangan karakter pada barisan X dengan spasi pada barisan

Y

$$\begin{array}{ccccccc} X_1 & X_2 & X_3 & \dots & X_i & \dots & X_n \\ Y_1 & Y_2 & Y_3 & \dots & _ & \dots & Y_m \end{array}$$

b. Pemasangan karakter pada barisan Y dengan spasi pada barisan

X

$$\begin{array}{ccccccc} X_1 & X_2 & X_3 & \dots & \boxed{-} & \dots & X_n \\ Y_1 & Y_2 & Y_3 & \dots & Y_j & \dots & Y_m \end{array}$$

2.2. Pemberian bobot pada penyejajaran dua barisan DNA

Telah dijelaskan sebelumnya bahwa penyejajaran DNA dilakukan untuk melihat seberapa besar tingkat kemiripan antar barisan. Besar tingkat kemiripan ini diukur dari suatu skor penyejajaran. Skor penyejajaran merupakan suatu besaran yang melambangkan tingkat kemiripan antar barisan DNA

Pada subbab ini akan dijelaskan mengenai pemberian skor pada penyejajaran DNA. Pada penyejajaran barisan dipasangkan masing-masing karakter pada suatu barisan dengan karakter atau spasi pada barisan lain. Setiap pemasangan tersebut akan diberikan sebuah skor pemasangan. Skor Penyejajaran merupakan total dari seluruh skor pemasangan antar karakter pada pasangan barisan. Skor pemasangan apabila dua karakter sama haruslah ditetapkan sebagai suatu nilai positif. Hal ini disebabkan kedua barisan dikatakan semakin mirip jika skor penyejajarannya tinggi, sementara skor penyejajaran kedua barisan tinggi apabila skor setiap pemasangan karakter yang sama bernilai tinggi. Sebaliknya, nilai ketidakcocokan

ditetapkan sebagai nilai negatif atau nol. Hal yang sama juga berlaku untuk pemasangan karakter dengan spasi.

Berikut contoh ketentuan pemberian skor pada tugas akhir ini, yaitu

1. pemasangan antarkarakter yang sama akan diberikan skor 2,
2. pemasangan antarkarakter yang berbeda akan diberikan skor -1,
3. pemasangan karakter dengan spasi akan diberikan skor -1.

- Pada pemasangan *barisan* X dengan barisan Y

Jika $x_i = y_j$, maka pemasangan tersebut diberi skor 2.

Jika $x_i \neq y_j$, maka pemasangan tersebut diberi skor -1.

- Pada pemasangan x_i dengan spasi dan pemasangan spasi dengan y_j diberikan skor -1. Begitu pula untuk semua pemasangan karakter dengan spasi diberikan skor -1.

Contoh :

Diberikan penyejajaran barisan barisan :

$X = G A G G T A$

$Y = G _ A T T _$

- Skor pemasangan karakter G pada barisan X dengan karakter G pada barisan Y adalah 2.

G	A	G	G	T	A
G	_	G	T	T	_

- Skor pemasangan karakter G pada barisan X dengan karakter T pada barisan Y adalah (-1).

G	A	G	G	T	A
G	_	G	T	T	_

- Skor pemasangan karakter A pada barisan X dengan spasi pada barisan Y adalah (-1)

G	A	G	G	T	A
G	_	G	T	T	_

2.3. Menentukan Penyejajaran yang Optimum

Penyejajaran yang optimum adalah yang memiliki total skor penyejajaran yang paling besar dari semua penyejajaran yang mungkin.

Contoh :

Diberikan dua barisan :

Barisan yang pertama adalah : T T G A

Barisan yang kedua adalah : T A A

Maka terdapat tiga kemungkinan penyejajaran skor penyejajaran terbesar yaitu :

1. TTGA
T_AA

Dengan skor penyejajaran yaitu $(2) + (-1) + (-1) + (2) = 2$

Atau

2. TTGA
TA_A

Dengan skor penyejajaran yaitu $(2) + (-1) + (-1) + (2) = 2$

Atau

3. TTGA
_TAA

Dengan skor penyejajaran yaitu $(-1) + (2) + (-1) + (2) = 2$

2.4. Penyejajaran Global dan Penyejajaran Lokal

Penyejajaran Global digunakan untuk melihat kemiripan dari dua barisan DNA secara keseluruhan

Contoh :



Segmen merupakan suatu subbarisan yang dibentuk dari unsur-unsur suatu barisan yang berurutan.

Contoh :

Diberikan barisan X = 'A A A G G G T T T C C C'

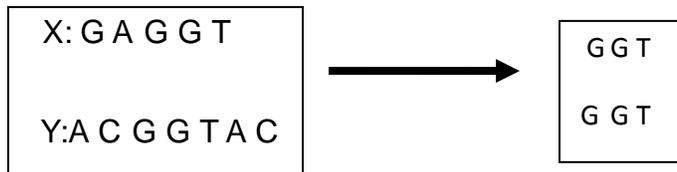
Barisan 'A A A G' merupakan segmen dari barisan X

Barisan 'G G T T T C' merupakan segmen dari barisan X

Barisan 'G A G A' bukan merupakan segmen dari barisan X

Penyejajaran lokal digunakan untuk mencari pasangan segmen dari dua barisan dengan kemiripan yang paling besar.

contoh penyejajaran lokal:



BAB III

METODE SMITH-WATERMAN

Metode Smith-Waterman bukan merupakan penyejajaran global melainkan penyejajaran lokal. Disini kita akan mencari pasangan segmen dari dua barisan yang memiliki tingkat kemiripan paling besar dibandingkan pasangan segmen yang lain.

Misalkan diberikan Barisan

$X = x_1x_2x_3\dots x_{n-1}x_n$ dengan panjang karakter n , dan barisan

$Y = y_1y_2y_3\dots y_{m-1}y_m$ dengan panjang karakter m .

Maka kita akan mencari pasangan segmen dari kedua barisan tersebut yang memiliki kemiripan paling besar

pertama-tama kita akan membuat tabel berukuran $n+1$ dan $m+1$

	0	X_1	X_2	X_{n-1}	X_n
0	$H_{0,0}$	$H_{0,1}$	$H_{0,2}$	$H_{0,n-1}$	$H_{0,n}$
y_1	$H_{1,0}$	$H_{1,1}$	$H_{1,2}$	$H_{1,n-1}$	$H_{1,n}$
y_2	$H_{2,0}$	$H_{2,1}$	$H_{2,2}$	$H_{2,n-1}$	$H_{2,n}$
:	:	:	:		:	:
:	:	:	:		:	:
Y_{m-1}	$H_{m-1,0}$	$H_{m-1,1}$	$H_{m-1,2}$	$H_{m-1,n-1}$	$H_{m-1,n}$
y_m	$H_{m,0}$	$H_{m,1}$	$H_{m,2}$	$H_{m,n-1}$	$H_{m,n}$

3.1. Menentukan Pembobotan pada Metode Smith-Waterman

Seperti yang dijelaskan sebelumnya pada penyejajaran barisan terdapat skor penyejajaran yang menentukan tingkat kemiripan dari dua barisan. Pada metode Smith-Waterman juga terdapat skor penyejajaran yang menentukan tingkat kesamaan dari dua segmen dari dua barisan. Skor penyejajaran pada metode ini juga merupakan total dari skor pemasangan antar karakter pada kedua segmen. Skor pemasangan yang digunakan yaitu:

Jika diberikan Barisan

$X = x_1x_2x_3\dots x_{n-1}x_n$ dengan panjang karakter n, dan barisan

$Y = y_1y_2y_3\dots y_{m-1}y_m$ dengan panjang karakter m.

Maka bobot pada pemasangan karakter x_i pada barisan X dengan y_j pada barisan Y, dimana $1 \leq i \leq n$ dan $1 \leq j \leq m$

$$S(x_i, y_j) = \begin{cases} 1.0 & \text{jika } x_i = y_j \\ -0.3 & \text{jika } x_i \neq y_j \end{cases}$$

W_k = bobot pemasangan karakter pada barisan dengan spasi yang memiliki panjang k

$$= 1.0 + 0.3 * k \text{ dimana } k \text{ adalah panjang spasi}$$

Contoh:

Pada penyejajaran berikut :

```
A G T G G T
_ _ A _ G A
```

Bobot dari pemasangan karakter T pada barisan pertama dengan A pada barisan kedua

A G T G G T
 _ _ A _ G A

Adalah -0,3

Dan bobot pemasangan karakter G pada barisan pertama dengan G pada barisan kedua

A G T G G T
 _ _ A _ G A

Adalah 1.0

Contoh dari pemasangan barisan dengan spasi

bobot dari penyejajaran

A G T G G T
 _ _ A _ G A

Adalah : $1.0 + 0,3 * 2 = 1 + 0,6 = 1,6$

3.2. Inisialisasi pada Metode Smith-Waterman

Pertama kita akan membuat tabel skor untuk metode Smith-Waterman.

	0	X ₁	X ₂	X _{n-1}	X _n
0	H _{0,0}	H _{0,1}	H _{0,2}	H _{0,n-1}	H _{0,n}
y ₁	H _{1,0}	H _{1,1}	H _{1,2}	H _{1,n-1}	H _{1,n}
y ₂	H _{2,0}	H _{2,1}	H _{2,2}	H _{2,n-1}	H _{2,n}
⋮	⋮	⋮	⋮		⋮	⋮
⋮	⋮	⋮	⋮		⋮	⋮
Y _{m-1}	H _{m-1,0}	H _{m-1,1}	H _{m-1,2}	H _{m-1,n-1}	H _{m-1,n}
y _m	H _{m,0}	H _{m,1}	H _{m,2}	H _{m,n-1}	H _{m,n}

Inisialisai pada metode Smith-Waterman

$H_{0,0}$ = skor maksimum sebelum adanya pemasangan karakter.

= 0

$H_{k,0}$ = pemasangan k karakter awal pada barisan X dengan spasi.

= 0

Karena pemasangan tersebut menandakan tidak adanya kesamaan pada pemasangan segmen k karakter awal pada barisan X dengan spasi sepanjang k pada barisan Y maka kita memberi skor pemasangannya

bernilai 0 sehingga pemasangan tersebut tidak mempengaruhi nilai dari skor penyejajaran segmen yang lain.

Contoh :

Pada penyejajaran barisan

X = A A A G G T

Y = _ _ _ G G T

Skor penyejajaran dari penyejajaran tersebut merupakan total skor dari skor sebelum adanya pemasangan karakter ditambah dengan skor pemasangan tiga karakter A padan barisan X dengan tiga spasi pada barisan Y ditambah skor pemasangan G dari barisan X dengan G dari barisan Y ditambah skor pemasangan karakter G dari barisan X dengan karakter G dari barisan Y ditambah skor pemasangan karakter T dari barisan X dengan karakter T dari barisan Y.

Sehingga skor penyejajarannya menjadi :

$$0 - (1 + 0,3 \times 3) + 1 + 1 + 1 = 0 - 1,9 + 3 = 1,1$$

sementara pasangan segmen yang memiliki kemiripan paling tinggi adalah

G G T

G G T

Dengan skor penyejajarannya adalah :

$$1 + 1 + 1 = 3$$

Agar skor penyejajaran dari pasangan segmen

G G T

G G T

yang merupakan bagian dari penyejajaran

A A A G G T

_ _ _ G G T

akan tetap bernilai 3 maka pasangan dari 3 karakter awal pada barisan tersebut yang berpasangan dengan spasi sepanjang 3 harus diberi nilai 0.

$H_{0,l}$ = Pemasangan l karakter awal pada barisan Y dengan spasi.

= 0

Similar dengan inialisasi dari $H_{k,0}$, karena pemasangan tersebut menandakan tidak adanya kesamaan pada pemasangan segmen k karakter awal pada barisan Y dengan spasi sepanjang l pada barisan X maka kita memberi skor pemasangannya bernilai 0 sehingga pemasangan tersebut tidak mempengaruhi nilai dari skor penyejajaran segmen yang lain.

$H_{i,j}$ = Skor maksimum dari penyejajaran dua segmen yang berakhir di x_i
dan y_j

Formula untuk $H_{i,j}$

$$H_{i,j} = \max \{ H_{i-1,j-1} + s(x_i, y_j), \max_{k \geq 1} (H_{i-k,j} - W_k), \max_{l \geq 1} (H_{i,j-l} - W_l), 0 \}$$

$$1 \leq i \leq n ; 1 \leq j \leq m$$

Formula tersebut terjadi dikarenakan pada penyejajaran yang berakhir pada x_i dari barisan X dan y_j dari barisan Y terdapat empat kemungkinan

1. Pemasangan terakhir dari penyejajaran tersebut adalah pemasangan karakter x_i dan karakter y_j sehingga skor penyejajarannya adalah skor penyejajaran yang berakhir pada x_{i-1} dan y_{j-1} ditambahkan dengan skor pemasangan x_i dengan y_j :

$$H_{i,j} = H_{i-1,j-1} + s(x_i, y_j)$$

2. Pemasangan terakhir dari penyejajaran tersebut adalah pemasangan k karakter pada barisan X dengan spasi sepanjang k pada barisan Y sehingga skor penyejajarannya adalah skor penyejajaran yang berakhir pada x_{i-k} dan y_j dikurang skor pemasangan k karakter dengan spasi.

$$H_{i,j} = (H_{i-k,j} - W_k)$$

3. Pemasangan terakhir dari penyejajaran tersebut adalah pemasangan l karakter pada barisan Y dengan spasi sepanjang l pada barisan X sehingga skor penyejajarannya adalah skor penyejajaran yang berakhir pada x_i dan y_{j-l} dikurang skor pemasangan l karakter dengan spasi.

$$H_{i,j} = (H_{i,j-l} - W_l)$$

4. Jika ketiga pilihan diatas tidak lebih besar dari nilai 0 maka penyejajaran segmen yang berakhir pada x_i dan y_j dinyatakan tidak memiliki kemiripan. Agar pasangan segmen tersebut tidak mempengaruhi skor penyejajaran segmen yang lain maka skor penyejajaran pasangan segmen tersebut diberi nilai 0.

$$H_{i,j} = 0$$

Ilustrasi pengisian bobot pada $H_{i,j}$

	0	y_1	y_2	...	y_{j-1}	y_j
0	$H_{0,0}$	$H_{0,1}$	$H_{0,2}$...	$H_{0,j-1}$	$H_{0,j}$
x_1	$H_{1,0}$	$H_{1,1}$	$H_{1,2}$...	$H_{1,j-1}$	$H_{1,j}$
x_2	$H_{2,0}$	$H_{2,1}$	$H_{2,2}$...	$H_{2,j-1}$	$H_{2,j}$
\vdots	\vdots	\vdots	\vdots		\vdots	\vdots
\vdots	\vdots	\vdots	\vdots		\vdots	\vdots
x_{i-1}	$H_{i-1,0}$	$H_{i-1,1}$	$H_{i-1,2}$...	$H_{i-1,j-1}$	$H_{i-1,j}$
x_i	$H_{i,0}$	$H_{i,1}$	$H_{i,2}$...	$H_{i,j-1}$	$H_{i,j}$

3.3. Penelusuran Kembali

Setelah pengisian tabel skor selesai hingga pemasangan karakter yang terakhir maka akan dilakukan penelusuran kembali. $H_{i,j}$ merupakan skor penyejajaran maksimum dari pemasangan dua segmen dari dua barisan sehingga semakin besar nilai $H_{i,j}$ maka kemiripan dari dua segmen tersebut semakin besar. Berdasarkan hal tersebut maka pasangan segmen yang memiliki kemiripan paling besar akan memiliki nilai $H_{i,j}$ terbesar dari semua nilai H yang ada. Misalkan terdapat pasangan segmen dengan kemiripan

terbesar memiliki nilai $H_{i,j}$ maka segmen tersebut berakhir pada x_i dan y_j sehingga penelusuran kembali dimulai dari $H_{i,j}$.

Setelah mengetahui bahwa pasangan segmen dengan kemiripan paling tinggi berakhir pada x_i dan y_j selanjutnya kita akan mencari dimana pasangan segmen tersebut dimulai. Pada penelusuran kembali kita memulai dari nilai $H_{i,j}$ yang paling besar. Penelusuran kembali ini dilakukan sampai kita menemukan $H_{i,j}$ yang bernilai 0 yang menandakan pasangan segmen tersebut dimulai. Karena $H_{i,j}$ yang bernilai 0 menandakan bahwa pasangan segmen yang berakhir pada nilai $H_{i,j}$ tersebut tidak memiliki kemiripan sehingga pasangan segmen tersebut tidak termasuk dalam pasangan segmen yang memiliki kemiripan terbesar karena akan membuat kemiripan dari pasangan segmen tersebut berkurang.

3.4. Contoh penyejajaran Smith-Waterman

Berikut ini adalah beberapa contoh dari penyejajaran lokal barisan DNA.

Contoh1 :

X : G A G G T

Y : A C G G T A C

$s(x_i, y_j)$ = bobot pemasangan x_i (karakter ke i dari barisan X) dengan y_j (karakter ke j dari barisan Y)

= 1.0 jika $x_i = y_j$

= -0.3 jika $x_i \neq y_j$

W_k = bobot pemasangan karakter pada barisan dengan spasi yang memiliki panjang k .

= $1.0 + 0.3 * k$ dimana k adalah panjang spasi

Pertama kita akan membuat matriks H

	0	G	A	G	G	T
0	$H_{0,0}$	$H_{0,1}$	$H_{0,2}$	$H_{0,3}$	$H_{0,4}$	$H_{0,5}$
A	$H_{1,0}$	$H_{1,1}$	$H_{1,2}$	$H_{1,3}$	$H_{1,4}$	$H_{1,5}$
C	$H_{2,0}$	$H_{2,1}$	$H_{2,2}$	$H_{2,3}$	$H_{2,4}$	$H_{2,5}$
G	$H_{3,0}$	$H_{3,1}$	$H_{3,2}$	$H_{3,3}$	$H_{3,4}$	$H_{3,5}$
G	$H_{4,0}$	$H_{4,1}$	$H_{4,2}$	$H_{4,3}$	$H_{4,4}$	$H_{4,5}$
T	$H_{5,0}$	$H_{5,1}$	$H_{5,2}$	$H_{5,3}$	$H_{5,4}$	$H_{5,5}$
A	$H_{6,0}$	$H_{6,1}$	$H_{6,2}$	$H_{6,3}$	$H_{6,4}$	$H_{6,5}$
C	$H_{7,0}$	$H_{7,1}$	$H_{7,2}$	$H_{7,3}$	$H_{7,4}$	$H_{7,5}$

Dimana

$H_{0,0}$ = maksimum kesamaan sebelum adanya pemasangan karakter.

$$= 0.0$$

$H_{i,j}$ = maksimum kesamaan dari dua segmen yang berakhir di x_i dan y_j

$$H_{i,j} = \max \{ H_{i-1,j-1} + s(x_i, y_j), \max_{k \geq 1} (H_{i-k,j} - W_k), \max_{l \geq 1} (H_{i,j-l} - W_l), 0 \}$$

$$1 \leq i \leq n ; 1 \leq j \leq m$$

$H_{k,0}$ = pemasangan k karakter awal pada barisan X dengan spasi.

$$= 0.0$$

$H_{0,l}$ = pemasangan l karakter awal pada barisan Y dengan spasi.

$$= 0.0$$

Formula untuk $H_{i,j}$

5. Jika x_i dan y_j saling berpasangan maka :

$$H_{i,j} = H_{i-1,j-1} + s(x_i, y_j)$$

6. Jika x_i berada pada akhir dari spasi dengan panjang k maka

$$H_{i,j} = (H_{i-k,j} - W_k)$$

7. Jika y_j berada pada akhir dari spasi dengan panjang l maka

$$H_{i,j} = (H_{i,j-l} - W_l)$$

8. $H_{i,j} = 0$; untuk mencegah perhitungan nilai negatif, menandakan tidak ada kesamaan antara pasangan segmen yang berakhir pada x_i dan y_j

$$\begin{aligned} H_{1,1} &= \max \{ H_{1-1,1-1} + s(x_1, y_1), \max(H_{1-1,1} - W_1), \max(H_{1,1-1} - W_1), 0.0 \} \\ &= \max\{ 0-0.3, \max(0-(1+0.3*1)), \max(0-(1+1/3*1)), 0.0 \} = 0.0 \end{aligned}$$

$$\begin{aligned} H_{2,1} &= \max\{ H_{1,0} + s(x_2, y_1), \max(H_{2-1,1} - W_1, H_{2-2,1} - W_2), \max(H_{1,1-1} - W_1), 0.0 \} \\ &= \max\{0.0-0.3, \max(0.0-(1+1.3*1), 0.0-(1+0.3*2)), \max(0.0-(1+0.3*1)), 0.0 \} \\ &= 0.0 \end{aligned}$$

$$\begin{aligned} H_{3,1} &= \max\{ H_{2,0} + s(x_3, y_1), \max(H_{3-1,1} - W_1, H_{3-2,1} - W_2, H_{3-3,1} - W_3), \\ &\quad \max(H_{1,1-1} - W_1), 0.0 \} \\ &= \max\{0.0 + 1.0, \max(0.0 - W_1, 0.0 - W_2, 0.0 - W_3), \max(0.0 - W_1), 0.0 \} = 1.0 \end{aligned}$$

$$\begin{aligned} H_{4,1} &= \max\{ \{ H_{3,0} + s(x_4, y_1), \max_{k \geq 1} (H_{4-k,1} - W_k), \max_{l \geq 1} (H_{4,1-l} - W_l), 0.0 \} \\ &= \max\{0.0 + 1.0, \max(1.0 - W_1, 0.0 - W_2, 0.0 - W_3, 0.0 - W_4), \\ &\quad \max(0.0 - W_1), 0.0 \} = 1.0 \end{aligned}$$

$$\begin{aligned} H_{5,1} &= \max\{ \{ H_{4,0} + s(x_5, y_1), \max_{k \geq 1} (H_{5-k,1} - W_k), \max_{l \geq 1} (H_{5,1-l} - W_l), 0.0 \} \\ &= \max\{0-0.3, \max(1.0 - W_1, 1.0 - W_2, 0.0 - W_3, 0.0 - W_4, 0.0 - W_5), \max(0.0 - W_1 \\ &\quad), 0.0 \} = 0.0 \end{aligned}$$

$$\begin{aligned} H_{6,1} &= \max\{ \{ H_{5,0} + s(x_6, y_1), \max_{k \geq 1} (H_{6-k,1} - W_k), \max_{l \geq 1} (H_{6,1-l} - W_l), 0.0 \} \\ &= \max\{0-0.3, \max(0.0 - W_1, 1.0 - W_2, 1.0 - W_3, 0.0 - W_4, 0.0 - W_5, 0.0 - W_6 \\ &\quad), \max(0.0 - W_1), 0.0 \} = 0.0 \end{aligned}$$

$$\begin{aligned}
H_{7,1} &= \max\{ H_{6,0} + s(x_7, y_1), \max_{k \geq 1}(H_{7-k,1} - W_k), \max_{l \geq 1}(H_{7,1-l} - W_l), 0.0 \} \\
&= \max\{0.0-0.3, \max(0.0- W_1, 0.0- W_2, 1.0- W_3, 1.0- W_4, 0.0- W_5, 0.0- W_6, 0.0- W_7, \\
&\quad), \max(0.0- W_1), 0.0\} = 0.0
\end{aligned}$$

$$\begin{aligned}
H_{1,2} &= \max\{ H_{0,1} + s(x_1, y_2), \max(H_{1-1,2} - W_1), \max(H_{1,2-1} - W_1, H_{1,2-2} - W_2), 0.0 \} \\
&= \max\{ 0.0 + 1.0, 1.0-(1+1/3*1), 0.0-(1+1/3*1), 0.0\} = 1.0
\end{aligned}$$

$$\begin{aligned}
H_{2,2} &= \max\{ H_{1,1} + s(x_2, y_2), \max(H_{2-1,2} - W_1, H_{2-2,2} - W_2), \\
&\quad \max(H_{2,2-1} - W_1, H_{2,2-2} - W_2), 0.0 \} \\
&= \max\{0.0-1/3, 1.0-(1+1/3*1), 0.0-(1+1/3*1), 0.0\} = 0.0
\end{aligned}$$

$$\begin{aligned}
H_{3,2} &= \max\{ H_{2,1} + s(x_3, y_2), \max_{k \geq 1}(H_{3-k,2} - W_k), \max_{l \geq 1}(H_{3,2-l} - W_l), 0.0 \} \\
&= \max\{0.0-1/3, 0.0-(1+1/3*1), 1.0-(1+1/3*1), 0.0\} = 0.0
\end{aligned}$$

$$\begin{aligned}
H_{4,2} &= \max\{ H_{3,1} + s(x_4, y_2), \max_{k \geq 1}(H_{4-k,2} - W_k), \max_{l \geq 1}(H_{4,2-l} - W_l), 0.0 \} \\
&= \max\{1.0-1/3, 0.0-(1+1/3*1), 1.0-(1+1/3*1), 0.0\} = 0.7
\end{aligned}$$

$$\begin{aligned}
H_{5,2} &= \max\{ H_{4,1} + s(x_5, y_2), \max_{k \geq 1}(H_{5-k,2} - W_k), \max_{l \geq 1}(H_{5,2-l} - W_l), 0.0 \} \\
&= \max\{1.0-1/3, 0.7-(1+1/3*1), 0.0-(1+1/3*1), 0.0\} = 0.7
\end{aligned}$$

$$\begin{aligned}
H_{6,2} &= \max\{ H_{5,1} + s(x_6, y_2), \max_{k \geq 1}(H_{6-k,2} - W_k), \max_{l \geq 1}(H_{6,2-l} - W_l), 0.0 \} \\
&= \max\{0.0+1.0, 0.7-(1+1/3*1), 0.0-(1+1/3*1), 0.0\} = 1.0
\end{aligned}$$

$$\begin{aligned}
H_{7,2} &= \max\{ H_{6,1} + s(x_7, y_2), \max_{k \geq 1}(H_{7-k,2} - W_k), \max_{l \geq 1}(H_{7,2-l} - W_l), 0.0 \} \\
&= \max\{0.0-1/3, 1.0-(1+1/3*1), 0.0-(1+1/3*1), 0.0\} = 0.0
\end{aligned}$$

$$\begin{aligned}
H_{1,3} &= \max\{ H_{0,2} + s(x_1, y_3), \max_{k \geq 1}(H_{1-k,3} - W_k), \max_{l \geq 1}(H_{1,3-l} - W_l), 0.0 \} \\
&= \max\{0.0-1/3, 0.0-(1+1/3*1), 1.0-(1+1/3*1), 0.0\} = 0.0
\end{aligned}$$

$$\begin{aligned}
 H_{2,3} &= \max\{ H_{1,2} + s(x_2, y_3), \max_{k \geq 1} (H_{2-k,3} - W_k), \max_{l \geq 1} (H_{2,3-l} - W_l), 0.0 \} \\
 &= \max\{ 1.0 - 1/3, 0.0 - (1 + 1/3 * 1), 0.0 - (1 + 1/3 * 1), 0.0 \} = 0.7
 \end{aligned}$$

$$\begin{aligned}
 H_{3,3} &= \max\{ H_{2,2} + s(x_3, y_3), \max_{k \geq 1} (H_{3-k,3} - W_k), \max_{l \geq 1} (H_{3,3-l} - W_l), 0.0 \} \\
 &= \max\{ 0.0 + 1.0, 0.7 - (1 + 1/3 * 1), 0.0 - (1 + 1/3 * 1), 0.0 \} = 1.0
 \end{aligned}$$

$$\begin{aligned}
 H_{4,3} &= \max\{ H_{3,2} + s(x_4, y_3), \max_{k \geq 1} (H_{4-k,3} - W_k), \max_{l \geq 1} (H_{4,3-l} - W_l), 0.0 \} \\
 &= \max\{ 0.0 + 1.0, 1.0 - (1 + 1/3 * 1), 0.7 - (1 + 1/3 * 1), 0.0 \} = 1.0
 \end{aligned}$$

$$\begin{aligned}
 H_{5,3} &= \max\{ H_{4,2} + s(x_5, y_3), \max_{k \geq 1} (H_{5-k,3} - W_k), \max_{l \geq 1} (H_{5,3-l} - W_l), 0.0 \} \\
 &= \max\{ 0.7 - 1/3, 1.0 - (1 + 1/3 * 1), 0.7 - (1 + 1/3 * 1), 0.0 \} = 0.3
 \end{aligned}$$

$$\begin{aligned}
 H_{6,3} &= \max\{ H_{5,2} + s(x_6, y_3), \max_{k \geq 1} (H_{6-k,3} - W_k), \max_{l \geq 1} (H_{6,3-l} - W_l), 0.0 \} \\
 &= \max\{ 0.7 - 1/3, 0.3 - (1 + 1/3 * 1), 1.0 - (1 + 1/3 * 1), 0.0 \} = 0.3
 \end{aligned}$$

$$\begin{aligned}
 H_{7,3} &= \max\{ H_{6,2} + s(x_7, y_3), \max_{k \geq 1} (H_{7-k,3} - W_k), \max_{l \geq 1} (H_{7,3-l} - W_l), 0.0 \} \\
 &= \max\{ 1.0 - 1/3, 0.3 - (1 + 1/3 * 1), 0.0 - (1 + 1/3 * 1), 0.0 \} = 0.7
 \end{aligned}$$

$$\begin{aligned}
 H_{1,4} &= \max\{ H_{0,3} + s(x_1, y_4), \max_{k \geq 1} (H_{1-k,4} - W_k), \max_{l \geq 1} (H_{1,4-l} - W_l), 0.0 \} \\
 &= \max\{ 0.0 - 1/3, 0.0 - (1 + 1/3 * 1), 0.0 - (1 + 1/3 * 1), 0.0 \} = 0.0
 \end{aligned}$$

$$\begin{aligned}
 H_{2,4} &= \max\{ H_{1,3} + s(x_2, y_4), \max_{k \geq 1} (H_{2-k,4} - W_k), \max_{l \geq 1} (H_{2,4-l} - W_l), 0.0 \} \\
 &= \max\{ 0.0 - 1/3, 0.0 - (1 + 1/3 * 1), 0.0 - (1 + 1/3 * 1), 0.0 \} = 0.0
 \end{aligned}$$

$$\begin{aligned}
 H_{3,4} &= \max\{ H_{2,3} + s(x_3, y_4), \max_{k \geq 1} (H_{3-k,4} - W_k), \max_{l \geq 1} (H_{3,4-l} - W_l), 0.0 \} \\
 &= \max\{ 0.7 + 1.0, 0.0 - (1 + 1/3 * 1), 1.0 - (1 + 1/3 * 1), 0.0 \} = 1.7
 \end{aligned}$$

$$\begin{aligned}
 H_{4,4} &= \max\{ H_{3,3} + s(x_4, y_4), \max_{k \geq 1} (H_{4-k,4} - W_k), \max_{l \geq 1} (H_{4,4-l} - W_l), 0.0 \} \\
 &= \max\{ 1.0 + 1.0, 1.7 - (1 + 1/3 * 1), 1.0 - (1 + 1/3 * 1), 0.0 \} = 1.0
 \end{aligned}$$

$$\begin{aligned}
 H_{5,4} &= \max\{ H_{4,3} + s(x_5, y_4), \max_{k \geq 1} (H_{5-k,4} - W_k), \max_{l \geq 1} (H_{5,4-l} - W_l), 0.0 \} \\
 &= \max\{ 1.0 - 1/3, 2.0 - (1 + 1/3 * 1), 0.3 - (1 + 1/3 * 1), 0.0 \} = 0.7
 \end{aligned}$$

$$\begin{aligned}
 H_{6,4} &= \max\{ H_{5,3} + s(x_6, y_4), \max_{k \geq 1} (H_{6-k,4} - W_k), \max_{l \geq 1} (H_{6,4-l} - W_l), 0.0 \} \\
 &= \max\{ 1.0 - 1/3, 2.0 - (1 + 1/3 * 1), 0.3 - (1 + 1/3 * 1), 0.0 \} = 0.7
 \end{aligned}$$

$$\begin{aligned}
 H_{7,4} &= \max\{ H_{6,3} + s(x_7, y_4), \max_{k \geq 1} (H_{7-k,4} - W_k), \max_{l \geq 1} (H_{7,4-l} - W_l), 0.0 \} \\
 &= \max\{ 1.0 - 1/3, 2.0 - (1 + 1/3 * 1), 0.3 - (1 + 1/3 * 1), 0.0 \} = 0.7
 \end{aligned}$$

$$\begin{aligned}
 H_{1,5} &= \max\{ H_{0,4} + s(x_1, y_5), \max_{k \geq 1} (H_{1-k,5} - W_k), \max_{l \geq 1} (H_{1,5-l} - W_l), 0.0 \} \\
 &= \max\{ 1.0 - 1/3, 2.0 - (1 + 1/3 * 1), 0.3 - (1 + 1/3 * 1), 0.0 \} = 0.7
 \end{aligned}$$

$$\begin{aligned}
 H_{2,5} &= \max\{ H_{1,4} + s(x_2, y_5), \max_{k \geq 1} (H_{2-k,5} - W_k), \max_{l \geq 1} (H_{2,5-l} - W_l), 0.0 \} \\
 &= \max\{ 1.0 - 1/3, 2.0 - (1 + 1/3 * 1), 0.3 - (1 + 1/3 * 1), 0.0 \} = 0.7
 \end{aligned}$$

$$\begin{aligned}
 H_{3,5} &= \max\{ H_{2,4} + s(x_3, y_5), \max_{k \geq 1} (H_{3-k,5} - W_k), \max_{l \geq 1} (H_{3,5-l} - W_l), 0.0 \} \\
 &= \max\{ 1.0 - 1/3, 2.0 - (1 + 1/3 * 1), 0.3 - (1 + 1/3 * 1), 0.0 \} = 0.7
 \end{aligned}$$

$$\begin{aligned}
 H_{4,5} &= \max\{ H_{3,4} + s(x_4, y_5), \max_{k \geq 1} (H_{4-k,5} - W_k), \max_{l \geq 1} (H_{4,5-l} - W_l), 0.0 \} \\
 &= \max\{ 1.0 - 1/3, 2.0 - (1 + 1/3 * 1), 0.3 - (1 + 1/3 * 1), 0.0 \} = 0.7
 \end{aligned}$$

$$\begin{aligned}
 H_{5,5} &= \max\{ H_{4,4} + s(x_5, y_5), \max_{k \geq 1} (H_{5-k,5} - W_k), \max_{l \geq 1} (H_{5,5-l} - W_l), 0.0 \} \\
 &= \max\{ 1.0 - 1/3, 2.0 - (1 + 1/3 * 1), 0.3 - (1 + 1/3 * 1), 0.0 \} = 0.7
 \end{aligned}$$

$$\begin{aligned}
 H_{6,5} &= \max\{ H_{5,4} + s(x_6, y_5), \max_{k \geq 1} (H_{6-k,5} - W_k), \max_{l \geq 1} (H_{6,5-l} - W_l), 0.0 \} \\
 &= \max\{ 1.0 - 1/3, 2.0 - (1 + 1/3 * 1), 0.3 - (1 + 1/3 * 1), 0.0 \} = 0.7
 \end{aligned}$$

$$\begin{aligned}
 H_{7,5} &= \max\{ H_{6,4} + s(x_7, y_5), \max_{k \geq 1} (H_{7-k,5} - W_k), \max_{l \geq 1} (H_{7,5-l} - W_l), 0.0 \} \\
 &= \max\{ 1.0 - 1/3, 2.0 - (1 + 1/3 * 1), 0.3 - (1 + 1/3 * 1), 0.0 \} = 0.7
 \end{aligned}$$

	0	G	A	G	G	T
0	0.0	0.0	0.0	0.0	0.0	0.0
A	0.0	0.0	1.0	0.0	0.0	0.0
C	0.0	0.0	0.0	0.7	0.0	0.0
G	0.0	1.0	0.0	1.0	1.7	0.4
G	0.0	1.0	0.7	1.0	2.0	1.4
T	0.0	0.0	0.7	0.4	0.7	3.0
A	0.0	0.0	1.0	0.4	0.4	1.7
C	0.0	0.0	0.0	0.7	0.1	1.4

Elemen H yang memiliki nilai yang paling besar adalah : $H_{5,5} = 3.0$

Lakukan penelusuran kembali

$H_{5,5}$ didapat dari $H_{4,4} + s(x_5, y_5) = 3.0$

$H_{4,4}$ didapat dari $H_{3,3} + s(x_4, y_4) = 2.0$

$H_{3,3}$ didapat dari $H_{2,2} + s(x_3, y_3) = 1.0$

Karena nilai $H_{2,2} = 0$ maka prosedur penelusuran kembali telah selesai.

Jadi segmen yang memiliki tingkat kemiripan paling besar adalah :

G G T

G G T

Contoh 2 :

Barisan pertama : 'G T A T C'

Barisan kedua : 'G T T C'

Yang pertama harus dilakukan adalah membuat tabel skor dari penyejajaran tersebut :

	0	G	T	T	C
0	$H_{0,0}$	$H_{0,1}$	$H_{0,2}$	$H_{0,3}$	$H_{0,4}$
G	$H_{1,0}$	$H_{1,1}$	$H_{1,2}$	$H_{1,3}$	$H_{1,4}$
T	$H_{2,0}$	$H_{2,1}$	$H_{2,2}$	$H_{2,3}$	$H_{2,4}$
A	$H_{3,0}$	$H_{3,1}$	$H_{3,2}$	$H_{3,3}$	$H_{3,4}$
T	$H_{4,0}$	$H_{4,1}$	$H_{4,2}$	$H_{4,3}$	$H_{4,4}$
C	$H_{5,0}$	$H_{5,1}$	$H_{5,2}$	$H_{5,3}$	$H_{5,4}$

Setelah itu kita akan mencari nilai $H_{i,j}$:

$$H_{0,0} = 0,0$$

$$H_{k,0} = 0,0$$

$$H_{0,l} = 0,0$$

$$H_{1,1} = \max \{ H_{1-1,1-1} + s(x_1, y_1), \max(H_{1-1,1} - W_1), \max(H_{1,1-1} - W_1), 0.0 \} = 1.0$$

$$H_{2,1} = \max \{ H_{1,0} + s(x_2, y_1), \max(H_{2-1,1} - W_1, H_{2-2,1} - W_2), \max(H_{1,1-1} - W_1), 0.0 \} = 0.0$$

$$H_{3,1} = \max \{ H_{2,0} + s(x_3, y_1), \max(H_{3-1,1} - W_1, H_{3-2,1} - W_2, H_{3-3,1} - W_3),$$

$$\max(H_{1,1-1} - W_1), 0.0 \}$$

$$= 0.0$$

$$H_{4,1} = \max \{ \{ H_{3,0} + s(x_4, y_1), \max_{k \geq 1} (H_{4-k,1} - W_k), \max_{l \geq 1} (H_{4,1-l} - W_l), 0.0 \} \} = 0,0$$

$$H_{5,1} = \max \{ \{ H_{4,0} + s(x_5, y_1), \max_{k \geq 1} (H_{5-k,1} - W_k), \max_{l \geq 1} (H_{5,1-l} - W_l), 0.0 \} \} = 0.0$$

$$H_{1,2} = \max \{ H_{0,1} + s(x_1, y_2), \max(H_{1-1,2} - W_1), \max(H_{1,2-1} - W_1, H_{1,2-2} - W_2), 0.0 \} = 0.0$$

$$H_{2,2} = \max \{ H_{1,1} + s(x_2, y_2), \max(H_{2-1,2} - W_1, H_{2-2,2} - W_2),$$

$$\max(H_{2,2-1} - W_1, H_{2,2-2} - W_2), 0.0 \}$$

$$= 2.0$$

$$H_{3,2} = \max \{ H_{2,1} + s(x_3, y_2), \max_{k \geq 1} (H_{3-k,2} - W_k), \max_{l \geq 1} (H_{3,2-l} - W_l), 0.0 \} = 0.7$$

$$H_{4,2} = \max \{ H_{3,1} + s(x_4, y_2), \max_{k \geq 1} (H_{4-k,2} - W_k), \max_{l \geq 1} (H_{4,2-l} - W_l), 0.0 \} = 1.0$$

$$H_{5,2} = \max \{ H_{4,1} + s(x_5, y_2), \max_{k \geq 1} (H_{5-k,2} - W_k), \max_{l \geq 1} (H_{5,2-l} - W_l), 0.0 \} = 0.1$$

$$H_{1,3} = \max \{ H_{0,2} + s(x_1, y_3), \max_{k \geq 1} (H_{1-k,3} - W_k), \max_{l \geq 1} (H_{1,3-l} - W_l), 0.0 \} = 0.0$$

$$H_{2,3} = \max \{ H_{1,2} + s(x_2, y_3), \max_{k \geq 1} (H_{2-k,3} - W_k), \max_{l \geq 1} (H_{2,3-l} - W_l), 0.0 \} = 1.0$$

$$H_{3,3} = \max \{ H_{2,2} + s(x_3, y_3), \max_{k \geq 1} (H_{3-k,3} - W_k), \max_{l \geq 1} (H_{3,3-l} - W_l), 0.0 \} = 1.7$$

$$H_{4,3} = \max \{ H_{3,2} + s(x_4, y_3), \max_{k \geq 1} (H_{4-k,3} - W_k), \max_{l \geq 1} (H_{4,3-l} - W_l), 0.0 \} = 1.7$$

$$H_{5,3} = \max \{ H_{4,2} + s(x_5, y_3), \max_{k \geq 1} (H_{5-k,3} - W_k), \max_{l \geq 1} (H_{5,3-l} - W_l), 0.0 \} = 0.7$$

$$H_{1,4} = \max \{ H_{0,3} + s(x_1, y_4), \max_{k \geq 1} (H_{1-k,4} - W_k), \max_{l \geq 1} (H_{1,4-l} - W_l), 0.0 \} = 0.0$$

$$H_{2,4} = \max \{ H_{1,3} + s(x_2, y_4), \max_{k \geq 1} (H_{2-k,4} - W_k), \max_{l \geq 1} (H_{2,4-l} - W_l), 0.0 \} = 0.4$$

$$H_{3,4} = \max \{ H_{2,3} + s(x_3, y_4), \max_{k \geq 1} (H_{3-k,4} - W_k), \max_{l \geq 1} (H_{3,4-l} - W_l), 0.0 \} = 0.7$$

$$H_{4,4} = \max \{ H_{3,3} + s(x_4, y_4), \max_{k \geq 1} (H_{4-k,4} - W_k), \max_{l \geq 1} (H_{4,4-l} - W_l), 0.0 \} = 1.4$$

$$H_{5,4} = \max \{ H_{4,3} + s(x_5, y_4), \max_{k \geq 1} (H_{5-k,4} - W_k), \max_{l \geq 1} (H_{5,4-l} - W_l), 0.0 \} = 2.7$$

Tabel skor dari penyejajaran diatas adalah :

	0	G	T	T	C
0	0	0	0	0	0
G	0	1	0	0	0
T	0	0	2	1	0.4
A	0	0	0.7	1.7	0.7
T	0	0	1	1.7	1.4
C	0	0	0.1	0.7	2.7

Nilai $H_{i,j}$ yang memiliki nilai terbesar adalah : $H_{4,4} = 4.0$

Lakukan penelusuran kembali

$$H_{5,4} \text{ didapat dari } H_{4,3} + s(x_4, y_4) = 2.7$$

$$H_{4,3} \text{ didapat dari } H_{3,2} + s(x_3, y_3) = 1.7$$

$$H_{3,2} \text{ didapat dari } H_{2,2} - W_1 = 0.7$$

$$H_{2,2} \text{ didapat dari } H_{1,1} + s(x_2, y_2) = 2.0$$

$$H_{1,1} \text{ didapat dari } H_{0,0} + s(x_1, y_1) = 1.0$$

Karena nilai $H_{0,0} = 0$ maka prosedur dari penelusuran kembali telah selesai.

Jadi segmen yang memiliki tingkat kemiripan paling besar adalah:

G T A T C

G T T C

Dengan penyejajaran yang memiliki nilai paling besar adalah:

G T A T C

G T _ T C

Dengan nilai skor maksimum adalah 2.7.

3.4. Penggunaan Program dinamik pada Metode Smith-Waterman

Program dinamik adalah prosedur matematika yang di desain untuk tujuan utama meningkatkan efisiensi computer untuk masalah program matematika dengan membagi-bagi masalah itu menjadi lebih kecil, dengan demikian komputasinya menjadi lebih mudah. Program dinamik memiliki ciri2x memecahkan masalah dalam stage-stage, dengan tiap stage melibatkan tepat satu variable optimisasi. Komputasi pada stage berbeda dihubungkan secara komputasi rekursif untuk menemukan hasil solusi optimal yang mungkin untuk seluruh masalah.

Konsep dasar yang merupakan karakteristik masalah pemrograman dinamik antara lain

1. Suatu masalah dibagi kedalam beberapa tahapan, dengan keputusan kebijakan diperlukan pada tiap tahapnya.
2. Tiap tahap mempunyai sejumlah keadaan yang saling berhubungan.
3. Pengaruh keputusan kebijakan di tiap tahap adalah merubah keadaan sekarang menjadi keadaan yang terhubung dengan tahap berikutnya.
4. Prosedur penyelesaiannya dirancang untuk menemukan keputusan

kebijakan dari keseluruhan masalah.

5. Saat keadaan sekarang, suatu keputusan kebijakan untuk tahap yang belum dilalui tidak bergantung dari kebijakan yang telah diambil pada tahap sebelumnya.
6. Prosedur penyelesaian ditandai dengan menemukan keputusan kebijakan pada tahap terakhir.
7. Adanya suatu hubungan rekursif.

Dalam pendekatan program dinamik masalah penyejajaran barisan merupakan pencarian jalur dengan skor terbesar dari keadaan awal sampai keadaan akhir.

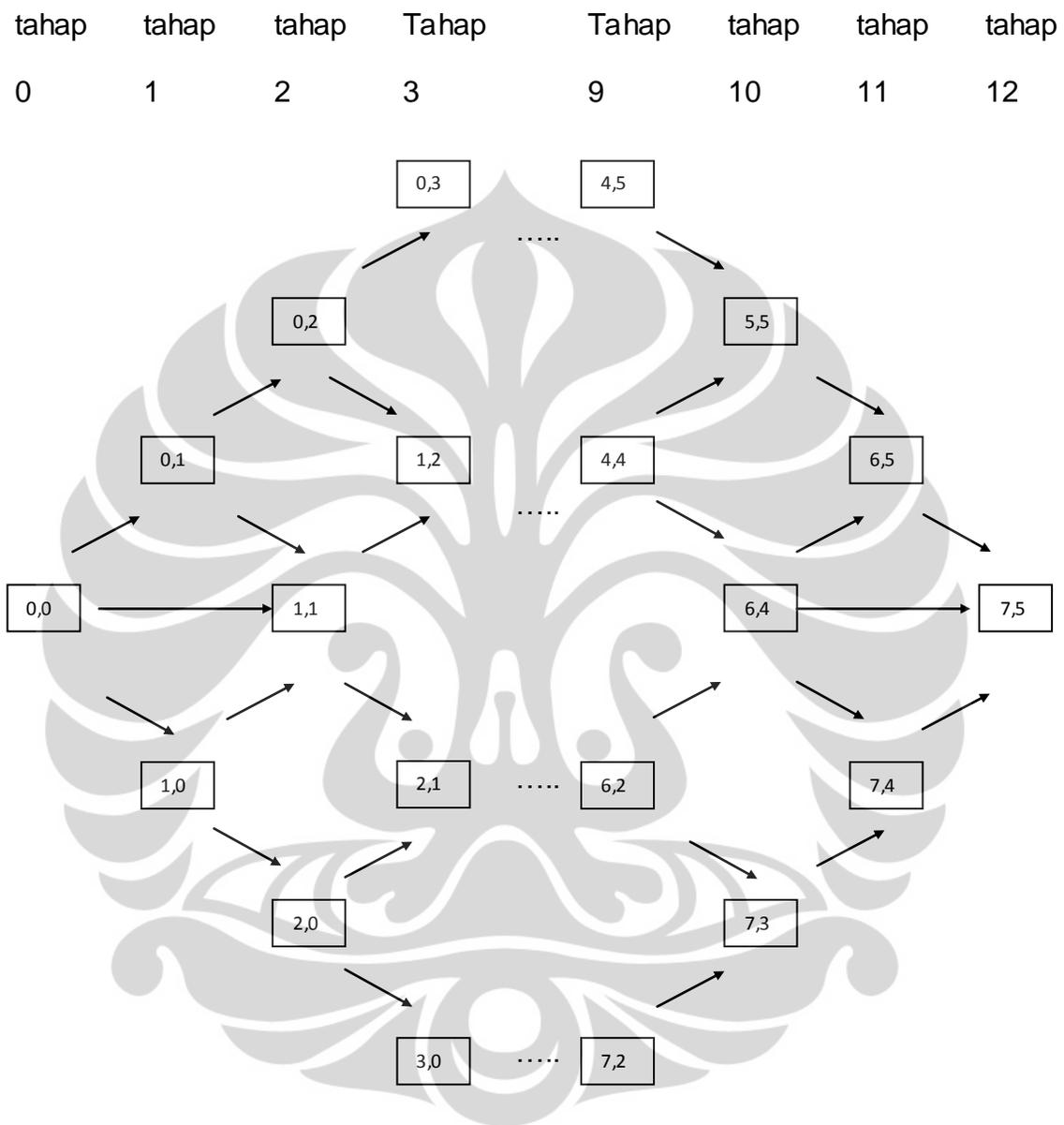
Berikut ini adalah contoh pendekatan program dinamik yang digunakan untuk menyelesaikan masalah dua barisan dengan metode Smith-Waterman.

Diberikan dua barisan

X : G A G G T

Y : A C G G T A C

Maka berikut ini adalah jaringan multistage pada masalah penyejajaran diatas :



Penjelasan pada gambar tersebut.

- Titik (0,0) adalah maksimum kesamaan sebelum adanya pemasangan karakter
 Skor maksimum di titik ini adalah 0,0 ($H_{0,0} = 0,0$)
- Titik (0,1), (0,2), ..., (0,n) pemasangan n karakter pada barisan X dengan spasi pada barisan Y
 Karena skor pemasangan karakter awal pada barisan X dengan spasi adalah 0
 Skor maksimum pada titik ini adalah 0,0 ($H_{0,1} = H_{0,2} = \dots = H_{0,5} = 0,0$)
- Titik (1,0), (2,0), (3,0), ..., (m,0) pemasangan spasi pada barisan X dengan m karakter pada barisan Y. karena skor pemasangan karakter awal pada barisan Y dengan spasi adalah 0. Maka skor maksimum pada titik $H_{1,0} = H_{2,0} = \dots = H_{7,0} = 0,0$
- Titik (1,1) adalah maksimum kesamaan yang berakhir di karakter pertama barisan X dan karakter pertama pada barisan Y. maka skor maksimum pada titik (1,1) berasal dari kemungkinan tiga skor titik yaitu (0,0), (1,0), dan (0,1). Maka dipilihlah titik yang memberikan skor maksimum pada titik (1,1) dan lebih besar dari 0. Karena skor maksimum dari ketiga titik tersebut nilainya lebih kecil dari 0 maka skor maksimum dari titik (1,1) adalah 0,0 ($H_{1,1} = 0,0$)

- Titik (1,2) adalah maksimum kesamaan yang berakhir di karakter kedua barisan X dan karakter pertama barisan Y. skor maksimum pada titik (1,2) bergantung pada skor di titik (0,2) , (1,1),(1,0),(0,1).). Maka dipilihlah titik yang memberikan skor maksimum pada titik (1,2) dan lebih besar dari 0. Kita mendapatkan bahwa skor maksimum pada titik (1,2) berasal dari titik (0,1) dengan skor maksimumnya yaitu 1,0. Maka nilai dari $H_{1,2}$ adalah 1,0
- dan seterusnya sampai pada titik (7,5).

Dari gambar tersebut proses pengisian dandan pemilihan skor maksimum di tiap tahap bergantung pada tahap sebelumnya

- tahap 0 merupakan inialisasi keadaan awal yaitu pada titik (0,0)
- tahap 1 terdiri dari dua keadaan yaitu titik (0,1) dan (1,0). Tahap ini bergantung pada tahap 0
- tahap 2 terdiri dari tiga keadaan yaitu titik (0,2),(1,1) dan (2,0) . tahap ini bergantung dari tahap 1 dan tahap 0.
- Tahap 3 terdiri dari empat keadaan yaitu titik(0,3),(1,2),(2,1),(3,). Tahap ini bergantung pada tahap 2, tahap 1 dan tahap 0.
- Dan seterusnya.

BAB IV

IMPLEMENTASI

Pada bab ini kita akan melihat implementasi dari algoritma dengan menggunakan software matlab 7. Input dari software tersebut adalah 2 barisan DNA dimana didalam DNA barisan itu terdiri dari empat karakter Alanine(A), Guanine(G),Thymine (T),dan Cytosine(C). Penulis mengambil barisan secara acak dari bioinformatic toolbox pada matlab 7. Sedangkan output-nya adalah hasil penyejajaran lokal yang optimal yang terdapat pada kedua barisan tersebut dengan skor maksimum. Berikut ini adalah hasil implementasinya.

4.1 Hasil Implementasi

Pada subbab ini akan dilihat hasil dari penyejajaran dari dua buah barisan untuk beberapa contoh barisan dengan panjang karakter yang berbeda-beda. Berikut ini adalah hasil percobaan yang dilakukan.

Contoh satu:

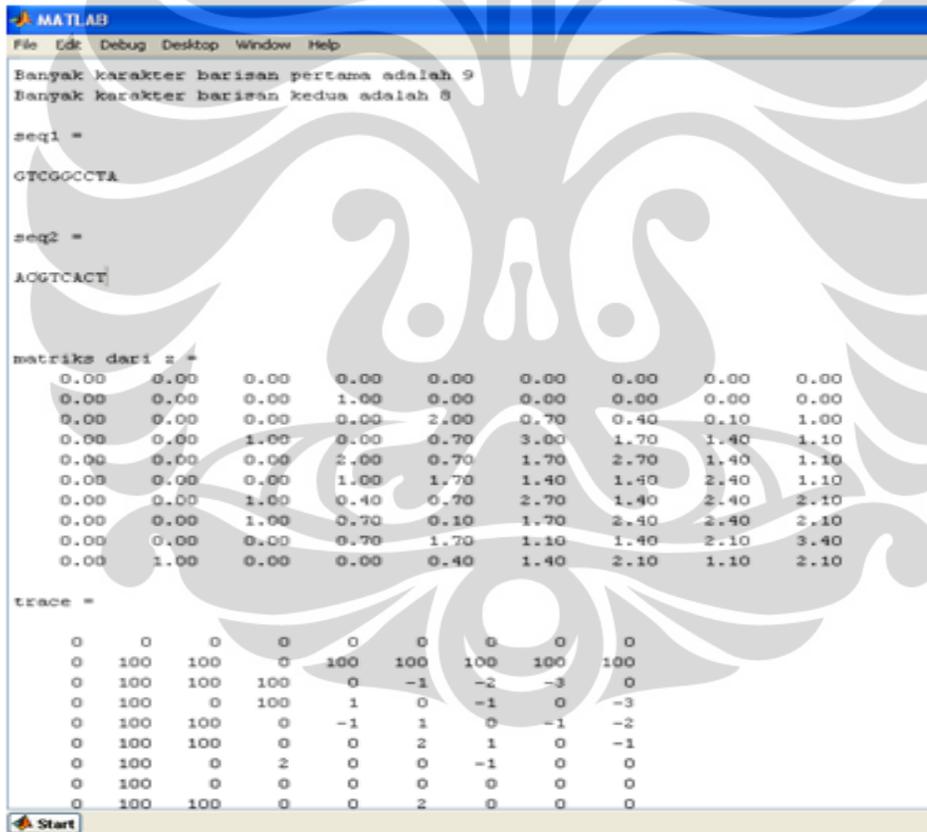
Barisan yang diambil secara acak dengan panjang karakter maksimal 10 karakter yang diacak menggunakan matlab 7.0.1 .

Didapat dua barisan sebagai berikut

Barisan 1 : 'G T C G G C C T A' dengan panjang karakter 9.

Barisan 2 : 'A C G T C A C T' dengan panjang karakter 8.

hasil output dari penyelarasan kedua barisan diatas dapat dilihat dari gambar berikut :



```

MATLAB
File Edit Debug Desktop Window Help

Banyak karakter barisan pertama adalah 9
Banyak karakter barisan kedua adalah 8

seq1 =
GTCGGCCTA

seq2 =
ACGTCACT

matriks dari z =
0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
0.00 0.00 0.00 1.00 0.00 0.00 0.00 0.00 0.00
0.00 0.00 0.00 0.00 2.00 0.70 0.40 0.10 1.00
0.00 0.00 1.00 0.00 0.70 3.00 1.70 1.40 1.10
0.00 0.00 0.00 2.00 0.70 1.70 2.70 1.40 1.10
0.00 0.00 0.00 1.00 1.70 1.40 1.40 2.40 1.10
0.00 0.00 1.00 0.40 0.70 2.70 1.40 2.40 2.10
0.00 0.00 1.00 0.70 0.10 1.70 2.40 2.40 2.10
0.00 0.00 0.00 0.70 1.70 1.10 1.40 2.10 3.40
0.00 1.00 0.00 0.00 0.40 1.40 2.10 1.10 2.10

trace =
0 0 0 0 0 0 0 0 0
0 100 100 0 100 100 100 100 100
0 100 100 100 0 -1 -2 -3 0
0 100 0 100 1 0 -1 0 -3
0 100 100 0 -1 1 0 -1 -2
0 100 100 0 0 2 1 0 -1
0 100 0 2 0 0 -1 0 0
0 100 0 0 0 0 0 0 0
0 100 100 0 0 2 0 0 0
  
```

```

MATLAB
File Edit Debug Desktop Window Help
0.00 0.00 0.00 1.00 1.70 1.40 1.40 2.40 1.10
0.00 0.00 1.00 0.40 0.70 2.70 1.40 2.40 2.10
0.00 0.00 1.00 0.70 0.10 1.70 2.40 2.40 2.10
0.00 0.00 0.00 0.70 1.70 1.10 1.40 2.10 3.40
0.00 1.00 0.00 0.00 0.40 1.40 2.10 1.10 2.10

trace =
0 0 0 0 0 0 0 0 0
0 100 100 0 100 100 100 100 100
0 100 100 100 0 -1 -2 -3 0
0 100 0 100 1 0 -1 0 -3
0 100 100 0 -1 1 0 -1 -2
0 100 100 0 0 2 1 0 -1
0 100 0 2 0 0 -1 0 0
0 100 0 0 0 0 0 0 0
0 100 100 0 0 2 0 0 0
0 0 100 100 0 0 0 0 1

nilai maksimum dari penyejajaran lokal 2 barisan dna ada pada z(9,9) adalah 3.40
sequence1 =
CGGC_CT
sequence2 =
CGTCACT
panjang dari penyejajaran lokal adalah 7
panjang segmen dari barisan 1 adalah 6
panjang segmen dari barisan 2 adalah 7
banyaknya pasangan yg sama adalah 5
banyaknya pasangan yg berbeda adalah 1
banyaknya deleksi adalah 1
>>

```

Matriks z pada gambar diatas menyatakan tabel skor dari penyejajaran lokal.

Pada gambar tersebut terlihat bahwa terdapat tabel skor dari penyejajaran berukuran 10×9 . Nilai maksimum dari tabel tersebut terdapat pada $z(9,9)$ yaitu 3,4. Setelah itu dilakukan penelusuran kembali dimulai dari $z(9,9)$.

Penelusuran kembali dilakukan dengan menggunakan matriks Trace. Pada matriks trace angka $k(1 \leq k \leq 10)$

pada $trace(i,j)$ menandakan bahwa x_i berada pada akhir dari spasi dengan panjang k . jika matriks $trace(i,j)$ bernilai $l(-10 \leq l \leq -1)$ menandakan bahwa y_j berada pada akhir dari spasi dengan panjang l . angka 100 pada $trace(i,j)$ menandakan bahwa $z(i,j)$ bernilai 0.

Pasangan segmen dari kedua barisan tersebut yang memiliki kemiripan maksimum adalah :

'C G G C C T' dari barisan pertama dan

'C G T C A C T' dari barisan kedua

penyejajaran lokal dari kedua barisan yang paling maksimum adalah :

C G G C _ C T

C G T C A C T

panjang dari penyejajaran lokal adalah 7

panjang segmen dari barisan pertama adalah 6

panjang segmen dari barisan kedua adalah 7

banyaknya pasangan yang sama dari penyejajaran lokal tersebut adalah 5

banyaknya pasangan yang berbeda adalah 1

banyaknya pasangan dengan spasi adalah 1

besar bobot maksimum dari penyejajaran tersebut adalah 3,4

contoh 2 :

barisan yang diambil secara acak dengan panjang karakter maksimal 10

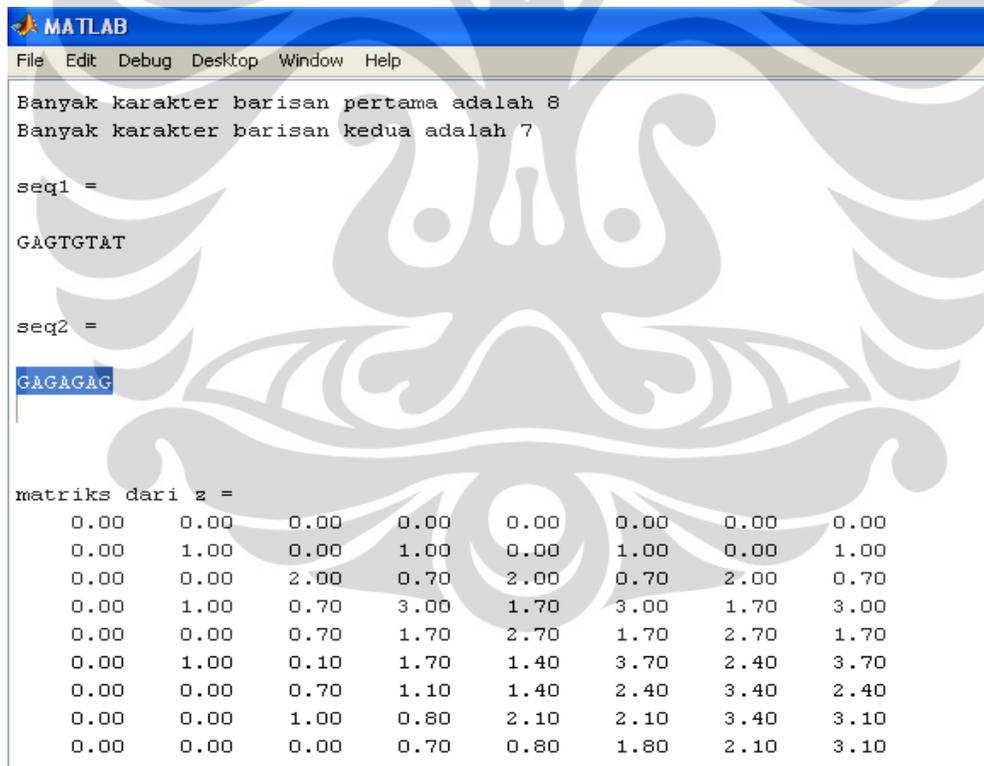
karakter

karakter yang diambil adalah :

barisan 1 : ' G A G T G T A T '

barisan 2 : ' G A G A G A G '

hasil output dari penyejajaran dapat dilihat dari gambar berikut :



```

MATLAB
File Edit Debug Desktop Window Help

Banyak karakter barisan pertama adalah 8
Banyak karakter barisan kedua adalah 7

seq1 =
GAGTGTAT

seq2 =
GAGAGAG

matriks dari z =
0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
0.00 1.00 0.00 1.00 0.00 1.00 0.00 1.00
0.00 0.00 2.00 0.70 2.00 0.70 2.00 0.70
0.00 1.00 0.70 3.00 1.70 3.00 1.70 3.00
0.00 0.00 0.70 1.70 2.70 1.70 2.70 1.70
0.00 1.00 0.10 1.70 1.40 3.70 2.40 3.70
0.00 0.00 0.70 1.10 1.40 2.40 3.40 2.40
0.00 0.00 1.00 0.80 2.10 2.10 3.40 3.10
0.00 0.00 0.00 0.70 0.80 1.80 2.10 3.10

```

```

trace =
  0   0   0   0   0   0   0   0
  0   0  100   0  100   0  100   0
  0  100   0  -1   0  -1   0  -1
  0   0   1   0  -1   0  -1   0
  0  100   0   1   0   1   0   1
  0   0   3   0   1   0  -1   0
  0  100   0   3   0   1   0   1
  0  100   0   4   0   2   0   0
  0  100  100   0   1   3   1   0

nilai maksimum dari pensejajaran lokal 2 barisan dna ada pada z(6,6) adalah 3.70

sequence1 =
GAGTG

sequence2 =
GAGAG

panjang dari pensejajaran lokal adalah 5
panjang segmen dari barisan 1 adalah 5
panjang segmen dari barisan 2 adalah 5
banyaknya pasangan yg sama adalah 4
banyaknya pasangan yg berbeda adalah 1
banyaknya delesi adalah 0
>>

```

Dari gambar tersebut terlihat tabel skor (tabel z) dari penyejajaran diatas berukuran 9 x 8. Nilai maksimum dari tabel tersebut terletak pada z(6,6) senilai 3,7 setelah itu dilakukan penelusuran kembali dari titik tersebut dengan menggunakan tabel trace.

Pasangan segmen dari kedua barisan yang memiliki kemiripan maksimum adalah :

G A G T G dari barisan pertama dan

G A G A G dari barisan kedua.

penyejajaran lokal dari kedua barisan yang paling maksimum adalah :

G A G T G

G A G A G

panjang dari penyejajaran lokal adalah 5

panjang segmen dari barisan pertama adalah 5

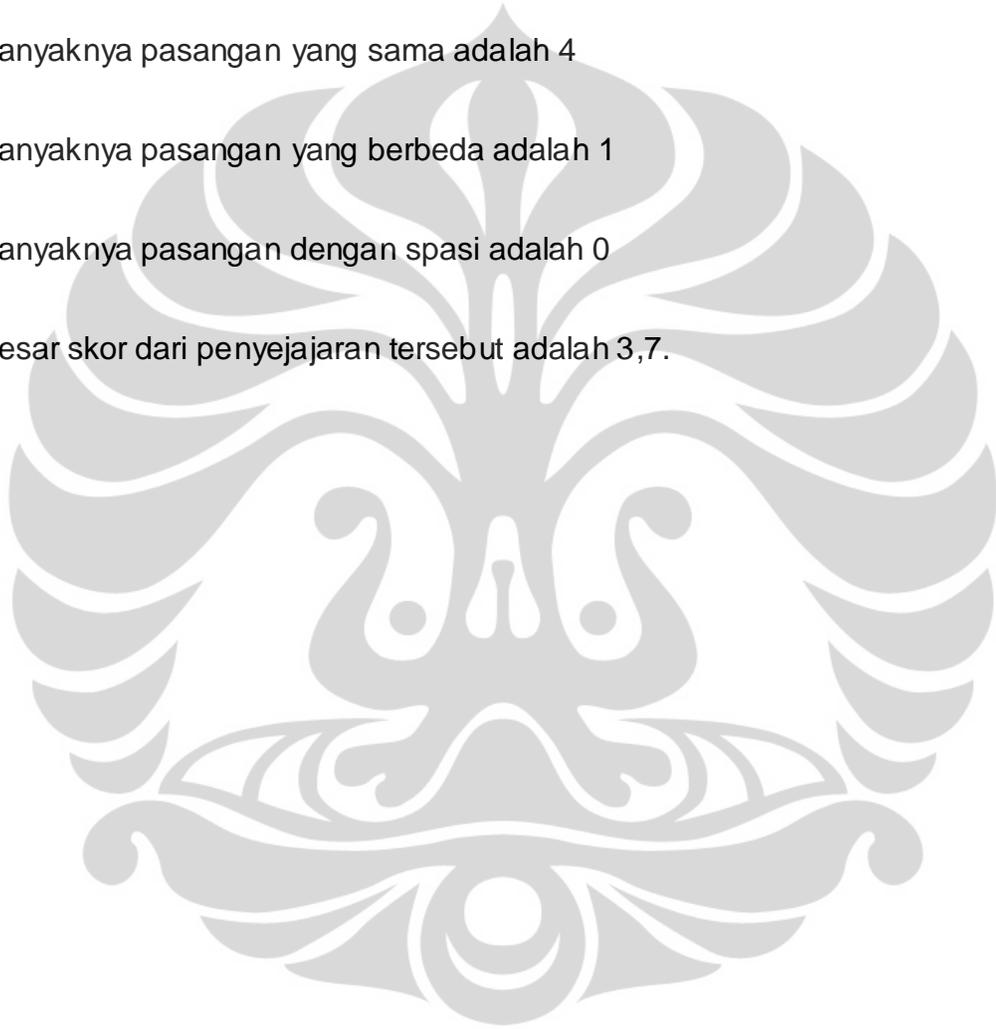
panjang segmen dari barisan kedua adalah 5

banyaknya pasangan yang sama adalah 4

banyaknya pasangan yang berbeda adalah 1

banyaknya pasangan dengan spasi adalah 0

besar skor dari penyejajaran tersebut adalah 3,7.



BAB V

KESIMPULAN

Berdasarkan dari percobaan-percobaan yang telah dilakukan kita diperoleh bahwa pada beberapa barisan-barisan tertentu, penyejajaran lokal dari barisan tersebut sama dengan penyejajaran global dari barisan itu.

Kelebihan pada metode Smith-Waterman adalah dapat menentukan pasangan segmen dari dua barisan yang memiliki kemiripan paling tinggi tanpa harus memperhitungkan seluruh kemungkinan pasangan segmen yang mungkin terjadi. Pada metode ini kita dapat pula mengurutkan pasangan segmen berdasarkan dari kemiripannya.

DAFTAR PUSTAKA

- [1] Isaev, Alexander. Introduction to Mathematical Method In Bioinformatics, Springer, 2006.
- [2] Jannah, Miftahul. Penyelesaian Pemadanan Barisan dengan Metode Needleman Wunsch pada penentuan kemiripan barisan Nucleotida.2006
- [3] Smith, T.F., Warterman, M.S., 1981, Identification Of Common Molecular Subsequences, J. Mol. Biol., 143, 195-197.

LAMPIRAN

Lampiran 1

```
clc
nseq1 = ceil(rand(1)*10);
nseq2 = ceil(rand(1)*10);
seq1 = randseq(nseq1, 'alphabet', 'dna');
fprintf('Banyak karakter barisan pertama adalah %d\n',nseq1);
seq2 = randseq(nseq2, 'alphabet', 'dna');
fprintf('Banyak karakter barisan kedua adalah %d\n',nseq2);
seq1
seq2
rseq1 = ['0',seq1];
rseq2 = ['0',seq2];

%pembuatan matriks z(bobot) dan trace(penelusuran kembali)
z = zeros(nseq1 + 1,nseq2 + 1);
trace = zeros(nseq1 + 1, nseq2 +1);

for i=2 : nseq1 + 1
    for j=2 : nseq2 + 1

        if(rseq1(i) == rseq2(j))
            p = z(i-1,j-1) + 1 ;
        else
            p = z(i-1,j-1) - (0.3);
        end

        deli = z(i-1,j) - (1 + 0.3);
        pdeli =1;
        for k = 2 : i - 1
            fakedeli = z(i-k,j) - (1 + (0.3)*k);
            if( fakedeli > deli)
                deli = fakedeli;
                pdeli = k;
            end
        end

        delj = z(i,j-1) - (1 + 0.3);
        pdelj = 1;
        for k = 2 : j - 1
            fakedelj = z(i,j-k) - (1 + (0.3)*k);
            if( fakedelj > delj)
                delj = fakedelj;
                pdelj = k;
            end
        end

        z(i,j) = max([p,deli,delj,0]);

        if (z(i,j)==p)
```

```

        trace(i,j) = 0;
    elseif(z(i,j) ==deli)
        trace(i,j) = pdeli;
    elseif(z(i,j) ==delj)
        trace(i,j)= -pdelj;
    elseif(z(i,j)==0)
        trace(i,j) = 100;
    end
end
end

fprintf('\n\n');
fprintf('matriks dari z = \n');

for i=1 : nseq1+1
    for j=1 : nseq2+1
        fprintf('\t%.2f',z(i,j));
    end
    fprintf('\n')
end

trace
%menentukan skor terbesar
[p,q] = max(z);
[mp,mq] = max(p);

fprintf('nilai maksimum dari penyejajaran lokal 2 barisan dna ada pada
z(%d,%d) adalah %.2f\n',q(mq),mq, z(q(mq),mq));

%penelusuran kembali
sequence1 = '';
sequence2 = '';
v = q(mq);
w = mq;
while( v > 1 && w > 1)
    if(trace(v,w) == 100)
        v=0;
        w=0;

    elseif(trace(v,w) == 0)
        sequence1=[rseq1(v), sequence1];
        sequence2=[rseq2(w), sequence2];
        v=v-1;
        w=w-1;

    elseif (trace(v,w) > 0)
        for i = 1 : trace(v,w)
            sequence1 = [rseq1(v),sequence1];
            sequence2 = ['_',sequence2];
        end
        v=v-trace(v,w);

    elseif (trace(v,w) < 0)

```

```

for j = 1 : ((-1) * trace(v,w))
    sequence1 = ['_', sequence1];
    sequence2 = [rseq2(w), sequence2];

end

w=w-((-1) * trace(v,w));

end
end

sequence1
sequence2
s = numel(sequence1);
fprintf('panjang dari penyejajaran lokal adalah %d\n',s);

%banyaknya k(kesamaan)m(tidak sama)d(spasi)pada penyejajaran
k=0;
m=0;
d=0;
for i=1 : s

    if(sequence1(i)=='_')
        d=d+1;
    elseif(sequence2(i)=='_')
        d=d+1;
    elseif(sequence1(i) == sequence2(i))
        k=k+1;
    elseif(sequence1(i) ~= sequence2(i))
        m = m+1;

    end

end

%panjang sequence 1 dan 2
pseq1 = 0;
pseq2 = 0;
for i=1 :s
    if(sequence1(i)~='_')
        pseq1 = pseq1+1;
    end
    if(sequence2(i)~='_')
        pseq2 = pseq2+1;
    end
end

fprintf('panjang segmen dari barisan 1 adalah %d\n',pseq1);
fprintf('panjang segmen dari barisan 2 adalah %d\n',pseq2);

fprintf('banyaknya pasangan yang sama adalah %d \n',k);
fprintf('banyaknya pasangan yang berbeda adalah %d \n',m);
fprintf('banyaknya spasi adalah %d\n',d);

```